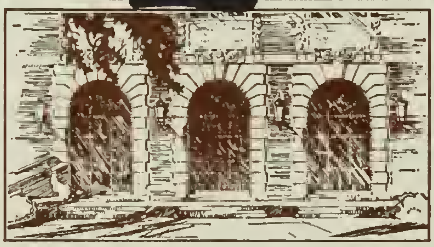


UIUCDCS-R-73-589

ARCHITECTURE OF A COMMUNITY MEDICAL HISTORY DATA BASE

by

Craig Alan Mills

August 7, 1973

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN · URBANA, ILLINOIS

ARCHITECTURE OF A COMMUNITY MEDICAL HISTORY DATA BASE

BY

CRAIG ALAN MILLS

B.S., University of Santa Clara, 1971

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1973

Urbana, Illinois

ACKNOWLEDGMENT

The author would like to express his appreciation to his advisor, Professor S. Ray, for his advice and criticism. The author would also like to express his gratitude to his wife, Linda, for translating his handwriting into a typed draft.

PREFACE

The medical profession has many cumbersome methods, such as most present medical codes, handwritten records, and the amount of paperwork involved in the care of a patient. These cumbersome methods and the establishment of Medicare, insurance forms, and other red tape which have become an intricate part of our society hinder the development of computer applications in the medical profession. This paper attempts to cover the many legal, security, and technical problems that arise. A general background of computer applications in the medical professions is given with emphasis on the usefulness of a medical history data base. From such computer applications, the medical profession will derive innumerable benefits. These benefits include more complete records, more individualized care, and efficient handling of paperwork.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Technology has created many new ways to cure people. However, with these new ways come new results, symptoms, questions, and answers that must be recorded. Politics have created Medicare and Medicaid and all the red tape that goes with it. In this day of rising prices, more people have more insurance creating more paperwork. Legal questions have prompted the use of more forms and more complete forms. With all of this additional information, the highly technological medical profession is still manually storing and retrieving information.

The hospital and private practices are businesses. However, they often lack the efficiency of their industrial counter parts even with industrial and financial leaders on the Board of Trustees. With the increase in cost and paperwork, it has become increasingly important for the hospital to be efficient, but being efficient does not necessarily mean less care for the patient. Efficiency could mean less cost to the patient and a more individualized service. Some hospitals have started to computerize and streamline their operations. One example is the State University Hospital of New York.

The State University Hospital of New York and Downstate Medical Center is unique in that it has had its activities controlled by an on-line computer system, Thomis (Total Hospital Operating and Medical Information System), since its opening in 1966. See in particular [22].

The patient's previous history is retrieved from disk. The computer then assigns a bed within the contraints of sex and locations, and preprinted, self-adhering labels for the patient are prepared by the computer. The nursing station is notified that the patient has arrived and the telephone office is informed which telephone is assigned to that patient.

The computer handles all transactions for both in-patients and out-patients although they require to some degree separate programs and files. In the case of a future order, the computer records the request and then informs the service area on the day the procedure is to be performed. Scheduling of procedures is done for the following service areas: Radiology, Physiology, Rehabilitation Medicine, Physical Therapy, Occupational Therapy, Neurology, Orthopedics, Plastic Surgery, Inhalation Therapy, Psychiatry, Opthalmalogy, and Operating Room.

When ordering from the pharmacy, at the same time a validation is printed out on the ordering terminal, a prescription label is printed containing all data legally necessary and a copy for the file is made. The stock inventory of the item ordered is reduced by the appropriate amount. If the inventory of the item falls below a specified level, the pharmacy is sent a notification to reorder. As is the case with all orders, the patient's bill is updated.

Thomis handles the reporting of quantitative laboratory results in the areas of clinical chemistry, coagulation, hematology, special hematology, serology, and clinical isotopes.

The results are entered by a clerk and then distributed by Thomis to the appropriate nursing stations. In-patient records are updated. Out-patient records are not usually available, so their results are summarized nightly and sent to the medical records department.

Thomis also includes an accounts receivable subsystem, management reports, extensive census reports, and statistical programs for research and activity analysis [22].

Even with the existence of such systems, computers are not being used to their full potential. Often a single hospital will attempt to computerize too much and cost, poor management, or lack of communication between users and installers will cause the collapse of the project. Other times, several small projects are completed but are not easily integrated into a harmonious unit.

One area that has not been successful is the storage and retrieval of a complete and detailed patient history data base. In this paper an attempt is made to design such a data base.

First, a description of what the computer can do to help hospitals is given in Chapter 2. Following this, in Chapter 3, the problems that will be incurred in developing information systems are briefly examined. Then, in detail the actual data base for a complete system is discussed with a view to the problems, expectations, and solutions that arise. Finally, in the conclusion, the main points are summarized and implementation requirements are outlined.

CHAPTER 2

SCENARIO OF A HOSPITAL INFORMATION SYSTEM

2.1 Hospital Information Systems

Like a business, hospitals also suffer wastes of time, energy, and money when computers are misapplied. For this reason, computers should be applied in modular fashion to areas with a well-defined purpose. At present, there are four major areas of hospital activity to which computers have successfully been applied. The first, since hospitals are very much a business, is data processing, where the most widespread use of computers in hospitals is in business systems. For example, business data processing, and in particular, data processing for accounts receivable, bills, and payroll are highly developed outside the hospital and easily adapted to any individual hospital.

The second and more challenging area is that of information retrieval. Here the peripheral display device gains a greater importance than the computer. It does not matter how economical a display device is, if it makes too much noise for a nursing station, for example, it does not belong in the hospital.

While airlines have had excellent information retrieval systems for many years, there are as yet no exceptional models in hospitals, although there are several working models which

will eventually be refined into a smooth functioning system. However, a major cause for frustration is the doctor who continues to compile medical records in a cumbersome way. Nevertheless, because of its sheer bulk, medical data will eventually be forced to be recorded in a structure that will permit computer manipulation.

The third area is management. Hospitals have management staffs and often trustee groups whose members are industrial, financial, and professional leaders who have had considerable management experience with computers and automation. Hence, the computer can be used for financial analysis and maintenance scheduling within the hospital itself.

The fourth area is the medical-clinical activity: diagnostic, research, and educational. The mini-computer is becoming very important in this area as the laboratory control element. The role of the mini-computer in the laboratory is threefold. Depending on the size of the laboratory computer and the interface with the hospital's central computer, the laboratory computer is expected to reduce turnaround of test results by assuming much of the clerical work. In some systems the computer will assign patient lab numbers, produce work lists, quality control results, specimen tube labels, patient records, tally sheets for lab workload analysis and charge lists for accounting, and route the output to the proper location. All this frees nurses and the technologist from clerical work and allows them to be better utilized.

Many laboratory procedures require computations to convert raw data obtained from analytical instruments into meaningful information. Computerization of these procedures, especially in an on-line real time system, reduces the chance of a mathematical or transcription error.

In some cases, a computer can monitor analytical instruments, make corrections for drift of baseline, interpret and chart repeated analysis of standards, decide when methods are out of control and when to repeat a method that appears in error and relay error messages to the technologists.

The focus of this paper will be on the second area, information retrieval. With the attempt to develop Hospital Information Systems (HIS) come several large problems. The problems are not only technical but legal, behavioral, social, and political. Two problems that should be mentioned in connection with HIS are data processing requirements for health insurance and the right of privacy for patients in regard to computerized data bases. Along with the diversity of forms required by the various insurance companies, Medicare and Medicaid are increasing substantially the amount of data processing for claims. The processing of the claims places a large burden on the hospital and in many cases forces them to computerize early.

Besides allowing the hospital to keep pace, the computer organizes and clarifies records. Before computerization patient records were often disorganized and frequently contained illegible handwritten records and notes. They were, consequently, implicitly safe from invasion of privacy. With computerization,

several problems will have to be solved before data bases are widely accepted. To assure continued patient confidence in the doctor, the computer must be covered by testimonial privilege. Also, the prevention of "leaks" must be considered, as there will be several programmers, technicians, nurses, etc. with access to the computer. Policies need to be developed on the use of medical information based on the purpose of the request, nature of the information being requested, individuals or organizations requesting the information, and need or lack of need for an informed written consent from the patient.

Hospital information can be divided into two categories: medical information and administrative information. However, the information in these two categories is not distinct. Both categories contain patient identification and a list of tests given to each patient, for the results in one category and for billing purposes in the other.

The most vital part of all the information is the patient identification. For instance, should the originating department request a test and the only legible identification they give is the patient's name, then the Lab must call the Information Center or Admitting Office to determine the location of the patient in order to transmit the results to the correct place and personnel. The carbon copy of the test request goes to the accounting department which organizes their records by patient numbers, so they too must inquire and waste time to gather the essential information.

In short, to make a Hospital Information System efficient, administrators and physicians must have a understanding of the information flow in the hospital: they must be aware of what information is needed by whom and when.

2.2 Community Medical History Data Bases

Previous discussion has shown that an in-house Hospital Information System is desirable, but desirability is not enough. Present cost of such a system would prove economically justifiable to only the largest institutions with over 1000 beds. Also, it has been estimated that nearly 70% of important clinical observations are recorded outside the hospital environment and at best only a rough sketch of the family medical history is available. The solution to these problems is to extend the Hospital Information System to include the entire community. This would provide a system not only economically feasible to small hospitals, but also to individual doctors. This Community Medical History Data Base would contain complete records gathered from all medical sources, including family histories. It is also possible that centralized billing and insurance processing is desirable and economical, but the remainder of this paper will concentrate on the development of the Community Medical History Data Base, forthwith referred to as the CMHDB. See in particular [21 and 47].

For the private doctor, CMHDB must be designed to compete with the present system. It should be just as quick for a nurse to retrieve information from CMHDB as it is presently

to find the hard copy. The nurse must be able to request some type of summary on the patient for the doctor to review before meeting the patient. During the examination the doctor may want to request more detail, or, in an emergency, pertinent information must be available immediately. All these conditions must be met before a CMHDB becomes viable to the individual doctors. However, not only must the requirements be met, they must be an improvement upon the existing system, especially in respect to retrieval time and completeness of records. For this a cathode ray display screen should be available for scanning the records for the needed information without producing a hard copy. This kind of response time requires an on-line system. By being on-line, the nurse does not have to worry if a requested file for a given appointment has arrived. In emergencies information is retrieved in seconds. With the short response time, hard copy information is minimized since more detail can be quickly retrieved.

With retrieval of data on-line, updating would most likely be on-line. The nurse of a particular doctor is more familiar with the terminology of that doctor than any central updating service would be. Also, the thought of 200-400 doctors sending incomplete forms to the central updating services makes it economically plausible that the updating should be done at each doctor's office. This also avoids double bookkeeping. Having the nurse do the updating requires CMHDB to contain sufficient safeguards to avoid accidental (or purposeful) removal

or changing of the wrong information. Posting the current date and other miscellaneous bookkeeping could be done automatically by CMHDB.

Some doctors or hospitals may want to use "Mumps" while others may establish their own language to cope with their particular terminology or special needs. Because of this, it is expected that each user could have his own customized interface. It will be up to the doctor or hospital to implement the interface. For further information on "Mumps", see [25].

CHAPTER 3

DATA BASE ARCHITECTURE

3.1 Legal Impediments

The changes required in the law to allow CMHDB to be efficient and effective poses a problem almost as immense as the development of CMHDB. There is no distinct body of law to be attacked. The government makes laws in response to influences, post de facto. As the law is presently stated, the creation of an automated system is precluded. See in particular [21 and 64].

Many states, such as California and Kansas, have regulations that allow automation of medical records after the original records have been manually created. In Maryland, as in other states, medical records are required to be manually created, i.e., legibly written by pen. Laws of this nature must be revised to allow efficient use of automation.

Regulations covering the contents of medical records vary from broad descriptive statements in Arizona, Iowa, and Ohio to Michigan's regulations which contain specific itemization of information to be included in the records. Generally, the medical record principles, standards, and interpretations promulgated by the Joint Commission on Accreditation of Hospitals meet the various state regulations. State courts have used them as evidence of the standard of care in negligence cases. Hence, any automated medical record system would do well to abide by these standards.

The Joint Commission with the interpretations under Standard II requires all medical records to contain:

1. Identification data and consent forms,
2. History of the patient,
3. Report of the physical examination,
4. Diagnostic and therapeutic orders,
5. Observations,
6. Reports of actions and findings, and
7. Conclusions.

The patient medical records, usually the most detailed, should include: patient's name, address, next of kin, and other pertinent data for identification and consents as deemed necessary by the hospital's administration and the medical staff.

The history of the patient should incorporate the chief complaints, details of past illness, inventory of systems, past history, social history, and family history. This history should be as unbiased and concise as humanly possible.

The report of the physical examination should contain all relevant findings resulting from an assessment of all the systems of the body. If a physical examination has been conducted recently, then only new information need be recorded with the understanding that no other information has changed since then.

Diagnostic and therapeutic orders written by authorized house staff members and other individuals who have assigned practice privileges, telephone orders written by licensed

nurses, and diet orders should be contained in the patient record.

Observation reports should include progress notes by the medical staff and house staff, consultation reports, nurses' notes, and entries by allied health personnel. Consultation reports should contain a written opinion by the consultant. Progress notes by the medical staff should give a chronological report of pertinent facts concerning the patient's recovery and should be sufficient to describe the changes in each of the patient's conditions and the results of the treatment. Nurses' notes and entries by allied health personnel should contain only meaningful information. Opinions requiring medical judgment should be authenticated only by authorized house staff members and those persons who have been assigned practice privileges.

Reports of actions and findings should include such items as reports of pathology and clinical laboratory examinations, radiology examinations, medical and surgical treatment, and any other diagnostic or therapeutic procedures. All diagnostic and therapeutic procedures should be recorded and authenticated in the medical record and may include any reports from out-of-hospital facilities. All treatment procedures performed must be documented in the medical record. The surgeon should record and sign a preoperative diagnosis prior to surgery. Operative notes, prepared after surgery, should contain a description of the findings, the techniques used, the tissue removed or altered, and post-operative diagnosis.

The conclusion includes provisional diagnosis, primary and secondary final diagnosis, clinical resume and necropsy reports. The provisional diagnosis should reflect the responsible practioner's evaluation of the patient's condition at the time of admission. All relevant discharge diagnoses should be recorded, using the terminology of a recognized system of disease nomenclature. The clinical resume should summarize the significant findings and events of the patient's hospitalization, his condition on discharge, and the recommendations and arrangements for future care. Where a necropsy is performed, provisional and anatomic diagnosis should be recorded on the medical record within 72 hours, where feasible, and the complete protocol should be made part of the record within three months.

A copy of the clinical resume is required to be sent to the medical practitioner or clinic responsible for subsequent care of the patient. This requirement would be implicitly met by CMHDB, since the resume is available to anyone who needs it once the clinical resume is entered.

Some comments should be made here concerning certain key words and phrases. The phrase "entries by allied health personnel," recognizes that persons other than medical or dental practitioners and licensed nurses may write entries. This is important because specialized personnel may be needed for data entry into CMHDB.

The words, "pertinent," "relevant," "meaningful" and "necessary," need to be defined to form uniform and standard records. It is generally accepted that whether or not automation is implemented medical records need to be reorganized.

Medical records at this time contain much useless information and are often missing useful information. Automation of medical records may prove to be the catalyst for this reorganization.

Then there are the key words "signature" and "authentication." The Joint Commission of Accreditation of Hospitals has been farsighted in recognizing the potential application of computer technology to medical records. This is seen in the Joint Commission definition of "authenticated": To prove authorship, for example, written signature, identifiable initials or computer key. Many states have regulations requiring handwritten signatures which will have to be modified before automated medical records can be effective.

Hospital medical records are retained anywhere from 7 to 25 years, depending on the state regulation. The American Hospital Association recommends 15 years. The Joint Commission interpretation states, "The length of time that records are to be kept is dependent upon the length of time that they may be needed for continuing patient care and for legal, research, or educational purposes." In any case, an automated medical record system would have to retain records for an indefinitely long period, which can be treated as permanent.

A more detailed discussion of the law and automated medical records can be found in [64].

3.2 Security

Access control to the data base has important implications in the content and structure of the data base. The

software security features will, of course, work in conjunction with hardware features. Nevertheless, software safety techniques have a dramatic effect on the data base. First of all, a list of eligible users must be maintained in a high level of the memory hierarchy. Accompanying every request must be a positive identification of the user. Sufficient identification information to recognize the user to a high confidence level plus the degree of access freedom the user has is information that will be frequently used.

It is possible that identification could be as simple as using a bank cash card where the user would enter a machine readable card and separately enter through a console a code. Identification could be further verified by having the computer asking suspicious users or randomly selected users such questions as birthdate, address, or questions from the user's medical file.

Access control is simplified if the data is structured in a modular fashion. This is so users restricted to a certain type of information can be easily given just that information. For instance, an orthodontist should not be allowed and it is unlikely that he would want information concerning the obstetrics of his patient. However, serious problems do arise as to whom is allowed what information. It would be difficult indeed to provide mutually exclusive data for the obstetrician and the orthopedist. This would imply that the data base should be organized into small blocks of data and that for each user there should be a list of those data blocks to which

the user has access. A solution to this problem is achieved by organizing the data base on a problem-oriented basis. The speciality of a doctor is defined by the problems he is trained to cure. A problem-oriented organization allows the doctor the information he needs while at the same time he can be restricted to just that information. Most hospitals presently organize their patient files by hospital stay, but each hospital stay usually has a list of problems in Hospital-International Classification of Diseases-Adapted (H-ICD-A) codes on the discharge summary sheet.

Most problem coding is either infeasible or at best inefficient for computer use. For example, Current Medical Information & Terminology (CMIT) codes spleen infraction as 05-3723 and myocardium infraction acute as 04-1921, and a new code for each infraction. This makes statistical searches difficult, to say the least, and lookup tables quite large. H-ICD-A and International Classification of Diseases-Adapted (ICD-A) are modifications of International Classification of Diseases (ICD) which was developed 100 years ago to record causes of death, and has a structure unsuitable for the computer.

In 1965, the College of American Pathologists published SNOP, Systemized Nomenclature of Pathology. It is presently being extended to become either the Lexicon of Medical Practice for Computer or Systematized Nomenclature of Medicine (SNOMed). SNOMed codes will contain five fields as follows:

1. <u>Topography</u> - anatomic sites of disease,

2. Morphology - names of visually apparent alterations of the cell, tissue or organ structure,

3. Etiology - a classification of causes of diseases and injuries,

4. Function - the terms for disturbances of physiologic function as in signs and symptoms,

5. Therapy and Diagnostic Procedures and Miscellaneous Hospital Functions - the names of forms of therapy and diagnostic procedures.

The first four fields are the same as SNOP and the fifth could be a modified version of the Current Procedural Terminology of the AMA.

The topography field consists of two parts, a site number and a sub-site number. There are 140 sites and a variable number of sub-sites, depending on the particular site. All 36 parts of the lung, for example, would be preceded by the site code 28. Greater detail will be required in the topography of SNOMed to satisfy the wide variety of specialists in the medical field.

The other three fields of SNOP have the same two part structure. The divisions of Morphology include:

1. Traumatic Abnormalities,

2. Congenital Malformations, Absences and Deformities,

3. Mechanical Abnormalities,

4. Inflammations and Repair,

5. Degenerations, Neerosis and Depositions,

6. Fine Structure and Cytologic Alterations,

7. Growth Alterations,

8. and 9. Neoplasms and Miscellaneous Conditions.

In Etiology the sections are:

1. Bacteria,

2. Rickettsiae,

3. Viruses,

4. Other Pathogenic Organisms,

5. Chemicals,

6. Reserved for Expansion,

7. and 8. Drugs,

9. Physical Agents of Injury.

It is hoped most microbiologic parts of this field will be contained in SNOMed.

The field of Function has:

1. Disorders of Elements, Ions, Simple Compounds, and Certain Metalloproteins,

2. Metabolic and Nutritional Disorders,

3. Enzyme Disorders,

4. Endoerine Disorders,

5. Blood Coagulation Disorders,

6. Immune Responses and Hypersensitivity Reactions,

7. Cardiovascular, Respiratory and Neuromuscular Disorders,

8. Mental, Psychic, and Personality Disorders,

9. Miscellaneous Functional Disorders.

There will be many additions to this field in SNOMed (more concerning SNOMed can be found in [69]).

With the code being stored in the data base, a central dictionary can be stored to translate the code into preferred terms and synonyms or just the preferred term with the

translation of preferred term into the local synonyms being done at the terminal. Reverse translation can be done similarly. In many cases, it may be quicker for the doctor or his nurse to deal directly with the SNOMed code. As an example, a visual aid for SNOP has been tested. Most terms are arranged according to topography. Their code numbers can be located in seconds by lifting the page to the proper site to select from the alphabetically arranged diagnoses. Such visual aids could be prepared for each of the different fields in medicine [69].

Using SNOMed for coding diagnoses provides quick coding of the diagnoses, easy manipulation and retrieval by computer, and most important of all a means of restricting the flow in information. Using this coding, users can be restricted in the kind of information they receive by storing a set of codes each user can request. Treating the coding as a five dimensional space, this restricting set would range from a list of points to a cross product of intervals on each axis. SNOMed also allows the user more freedom to request very specifically which disease he wishes information on by entering one number for one field to find all diseases with that characteristic in the patient's history. This also allows statisticians a wide latitude. Searches for certain combinations are much easier to locate.

Besides restricting the information requested to that information relevant to the user, a user can be restricted to read-only or write-only. A private doctor and his nurse, for example, would have access only to those records relevant to

his practice and then only through his terminal. It is possible that the doctor, owing to his reputation for legibility and completeness, may be allowed read-only access while the nurse might only have write-only access.

3.3 Directories

After solving the legal problems and deciding on security information, medical codes, and data base contents, the architecture of the data base can be discussed in depth. The first structure encountered are the directories.

Two main directories are needed in the data base, one for user identification and one for patient sub-directories. The user directory first must contain the user identification number and if a pass card system is used the code is needed. It is possible that these two numbers could be replaced by a digitized voice print or finger print, but these would require much more storage. Also needed is a list of terminals the user is permitted or expected to use. Should user identification number, code and terminal identification not match those in the directory, the computer may wish to question the user further to confirm identification or report the attempted misuse. For this the user's data base or patient identification number is needed. The user's patient file would contain his name, occupation and address and this information should not be duplicated in the user directory. Once the user has achieved clearance, his request must be checked to see if it is within the restricting set; that is, the information requested is relevant to his specialty. Whether the user has

read-only, write-only, or read and write access must also be determined. This user identification directory does not readily lend itself to compaction. The following illustrates the suggested contents for the directory with their approximate size in core.

User identification number	2 bytes
Code	2 bytes
Data base number	3 bytes
Flags (count of terminal numbers and size of restriction set)	2 bytes
Read-write Access	2 bits
Terminal numbers (variable)	1 byte each
Restriction set (variable)	10 byte minimum
Minimum number of bytes per user	20 bytes

Read-write access and terminal numbers might be combined into one number indicating a set of terminals the user can use and his read-write access. The restriction set will need overhead storage to indicate the number of subsets and whether each subset consists of a single point or a set of points (Figure 1).

If one allows for the possibility of 500 users, the user directory would require 10-15K bytes. This size directory could easily be contained in core on a medium-large computer. It is not expected that there will be a very dynamic turnover of users or very many changes in user access rights. Probably the most that need be expected is a weekly update. Because of this relative stability and the core contained size,

User ID#	Code	Patient ID#	a	b	c	d
e	Restriction Set					

a) Number of terminals (4 bits)

b) Size of restriction set (10 bits)

c) User read-write access rights (2 bits)

d) & e) Terminal identification numbers (1 byte each)

Figure 1. User directory element.

a rather fixed structure allowing rapid access can be used. Some permutations of the access code can be used for a simple table lookup of the particular user entry address. The user entry can then be accessed with this address for verifying identification.

Once clearance for the user and his request has been accomplished, the data base is ready to be entered. If the request is for the entering of new information or a new patient, then it is assumed the computer will automatically store the information and make all the needed entries in the appropriate tables. The "appropriate tables" just referred to will be discussed as viewed with respect to requests for information retrieval. Every request will contain a patient identification number and at least one SNOP number. Other information in the request might include dates indicating how recent or old the information should be, combination of SNOP numbers indicating a search for complications of a disease or request for specific test results.

Location of the patient directory is the first problem in answering a request. The patient identification number will be used to locate a patient directory, since names are not unique and phonetically similar names are easily misspelled. Once the particular directory is located, name, address, and other identification information can be sent back to the terminal for verification of the patient identification while the requested information is being retrieved. If the data base is designed to handle 100,000 patient histories, any table lookup

runs into storage problems. Using a transformation of the patient identification number to address the table and allowing three bytes for the directory address and directory length, 300K bytes of storage are needed for the table. Since this table is used with every request, small access times are needed. To avoid frequently retrieving sections of the table from secondary memory, as large a section as possible should be maintained in primary storage. Ideally the whole table would be in core. All of which means a large core is required.

Once the particular directory address has been retrieved from the table, the patient directory can be retrieved. Immediately upon retrieval identification information can be sent for verification of the patient number. The identification information should consist of name, current address, marital status, birth date, and physical description. The physical description would include height, weight, color of eyes and hair, sex, and race. This is sufficient information for the requestor to make a positive identification on the patient even if some of the information is wrong. By letting the requestor make the visual check, it saves the requestor from typing this information into the computer, saves the computer from making the match and deciding if sufficient information matches to make a positive identification, and finally occupies the requestor while the computer retrieves the needed information.

To retrieve the requested information, the computer would search a linked list of SNOMed codes in the patient directory. The list would include codes for family history, personality

history, and the social history, besides those for the medical history.

Because there are expected to be as many as 100,000 patient directories, storage conservation and data compaction are of prime importance. Many data compaction techniques are presented in the literature. The methods referred to in the following discussion are explained in detail in the references [45, 55, and 61]. The header information of the patient directories can be broken into three fields; name, address, and description. The name field would have the format: last, first, middle initial. As can be seen, using "b̲" and "#" to represent a blank and end of field respectively, coding ",b̲" and ".#" as one character each is feasible. Similarly, in a study whose results are presented in Figure 2, 10 other pairs occurred with sufficient frequency to warrant coding with a special character. With the code presented in Figure 2, an average of 4.14 bits were needed to encode one character. If a different format were used or the whole middle name was included, a new frequency study would have to be made. The codes are constructed such that from the first four bits of the code, the number of bits in the code can be determined. For example, all seven bit codes and only seven bit codes begin with "1111." Also, since the characters are ordered by frequency of occurrence, it can be concluded that four bit codes occur nearly 40 per cent of the time. Therefore, 40 per cent of the time a direct one step table lookup is all that is needed for the translation.

Character	Frequency	Probability	Code	Character	Frequency	Probability	Code
↗		.08726	0000	J	4,669	.01344	1111110
,b	21,654	.06231	0001	ON	4,204	.01210	1111111
E	21,572	.06208	0010	IL	4,005	.01152	0000
S	17,718	.05099	0011	K	3,927	.01130	0001
A	17,381	.05002	0100	F	3,412	.00982	0010
b̲	16,914	.04867	0101	EL	3,372	.00971	0011
I	14,523	.04180	0110	OR	3,227	.00928	0100
#	14,309	.04117	0111	EN	3,084	.00888	0101
M	13,990	.04026	1000	V	2,666	.00767	0110
L	13,779	.03966	10010	-	2,381	.00685	0111
C	12,078	.03475	10011	Z	2,002	.00576	1000
O	11,856	.03412	10100	Ɏ	1,000	.00288	10010
T	11,853	.03411	10101	X	499	.00143	10011
ER	10,967	.03156	10110	7	123	.00035	10100
R	10,860	.03125	10111	Q	117	.00033	10101
H	10,747	.03093	11000	Ø	102	.00029	10110
D	10,318	.02969	11001	1	97	.00028	10111
·#	9,701	.02792	11010	9	51	.00015	11000
B	7,532	.02167	11011	6	24	.00007	11001
P	7,319	.02107	111000	,	221	.00064	11010
G	7,191	.02069	111001	3	19	.00005	11011
AR	7,014	.02018	111010	5	16	---	111000
W	6,122	.01762	111011	4	15		111001
IN	5,983	.01721	1111000	2	15		111010
U	5,768	.01660	1111001	8	6		111011
AN	5,626	.01619	1111010	&	1		1111000
Y	5,414	.01558	1111011	•	-		1111001
N	5,203	.01498	1111100				
AL	4,850	.01395	1111101	Totals	347,497	.99969	

Figure 2. Source alphabet and codes used for name fields.

The address field usually has a format similar to:

NNN street name St. (1)

City, State NNNNN (2)

where "N" represents a decimal digit between 0 and 9. Line (1) can be coded with two codes, a prefix code and a suffix code. Different codes are needed because a significantly different distribution is found in line (1) than in the name field or in a general English text. The prefix code would be used for the sequences of numbers of ten followed by a N, E, S, or W. The suffix code would code the rest using a single character for such common combinations as "bst". Such codes are presented in Figures 3 and 4. Line (1) can be coded on the average with 3.79 bits per character, while line (2) can be compressed in several different ways, depending on the number of cities involved. With very few cities involved and one city of particular high frequency, a Huffman code can be used where 0 is assigned to the most frequent city-state combination, 10 to the second most frequent, 110 to the third, etc. With several cities a six or seven bit code could be used for a table lookup. With many cities, the states could be numbered for a table lookup and new statistical variable codes could be developed for the cities and zip codes.

The description field would be fixed length. Marital status can be represented by two bits (single, married, separated, divorced). Birth dates can be coded in 16 bits (5 for the day, 4 for the month, and 7 for the last two digits of the year). There are still only two sexes for 1 bit. Physical

Character	Frequency	Probability	Code
$\underline{b}$	29,500	.1965	000
↗	24,010	.1600	001
1	14,693	.0980	010
2	12,945	.0863	0110
3	9,196	.0612	0111
•	9,000	.0600	1000
5	7,907	.0527	1001
0	7,639	.0508	1010
4	7,319	.0487	1011
6	5,858	.0397	11000
7	4,607	.0307	11001
8	4,229	.0282	11010
N	4,000	.0267	11011
9	3,534	.0235	11100
W	1,500	.0100	11101
S	1,500	.0100	11110
E	1,500	.0100	11111
Totals	148,937	.9930	

Figure 3. Prefix codes used in address fields.

Character	Frequency	Probability	Code
E	18,824	.0858	0000
R	15,931	.0726	0001
T	15,866	.0723	0010
N	14,489	.0660	0011
A	13,633	.0622	0100
bST.#	12,500	.0570	0101
b	12,220	.0557	0110
O	11,898	.0542	0111
S	11,196	.0510	1000
L	9,875	.0449	10010
#	8,510	.0387	10011
D	8,416	.0383	10100
↗		.0360	10101
I	7,748	.0352	10110
H	7,397	.0337	10111
C	5,021	.0228	11000
M	4,256	.0194	11001
W	3,727	.0170	11010
G	3,509	.0160	11011
U	3,449	.0157	111000
P	3,278	.0149	111001
B	3,149	.0143	111010
bAVE.#	3,000	.0137	111011
K	2,601	.0118	1111000
.	2,430	.0111	1111001
Y	2,354	.0107	1111010
F	2,235	.0102	1111011
V	1,702	.0078	1111100
'	974	.0044	1111101
0	800	.0036	1111110
1	800	.0036	1111111
2	800	.0037	0000
3	800	.0036	0001
4	800	.0036	0010
5	800	.0037	0011
6	800	.0036	0100
7	800	.0036	0101
8	800	.0037	0110
9	800	.0036	0111
J	441	.0020	1000
Z	410	.0019	10010
-	257	.0012	10011
X	249	.0011	10100
Q	121	.0005	10101
@	42	.0002	10110
&	00	.0000	10111
Totals	219,708	1.0006	

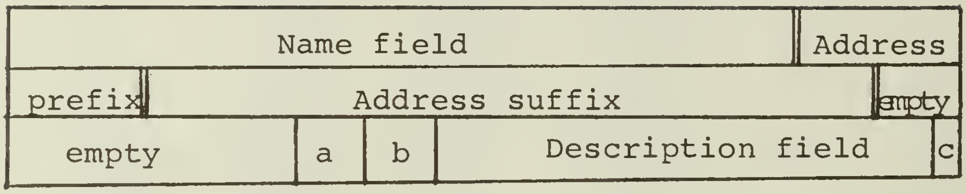
Figure 4. Suffix codes--used for address fields.

description can be coded as follows:

1.	Color of hair (black, brown, blonde, grey, others)	3 bits
2.	Color of eyes (blue, brown, others)	2 bits
3.	Race (Caucasian, Black, Chicano, Oriental, others)	3 bits
4.	Height (0-128 inches)	7 bits
	(0-256 centimeters)	8 bits
5.	Weight (0-512 pounds)	9 bits
	(0.0-204.8 kg.)	11 bits

Using inches and pounds the description field is 43 bits. Assuming 80 per cent of the time that the name field is less than or equal to twenty characters, line (1) is then less than or equal to twenty-five characters and line (2) can be represented by a seven bit code. Thus, to encode line (1), line (2), and the description field, a thirty byte record can be used with a high confidence level that a trailer will not be needed. If a trailer of an additional ten bytes is needed for the overflow, the last bit of the record can be used as a flag.

Once the header record has been retrieved, translated, and sent, the computer can search for the SNOMed codes requested. Each patient directory would hopefully contain at least three codes, those for the family, social, and personal histories. It is not expected, though, that there will be enough codes, on the average, to warrant a complex tree structure. A simple linear linked list with the codes in numerical order is expected to suffice. The five field SNOMed codes are ten bytes long. Any attempt to organize the codes would result in a high overhead



a) City-zipcode table address (6 bits)

b) State table address (6 bits)

c) Trailer flag (1 bit)

Note: a double vertical line represents an end of field character.

Figure 5. A patient directory information header.

in terms of storage, maintenance time, and search time. For instance, if each field is linked in a numerically ordered list, one additional byte for each field for the link address would be needed. Five more bytes for each of 20 codes for 100,000 patients is 10M bytes more for the patient directories and five linked lists to be altered each time a new code is entered. Each element of the list would contain a ten byte code, one byte for the number of references and five bytes for each reference consisting of two bytes for date and three bytes for address. Since two or more SNOMed codes could reference the same address, with some coding a reference might just point to another reference that has the same date and address at a savings of four bytes. If the list were ordered by the topography field, then a request for any history of inflammations within the last year would search every Morphology field in the list for code 4 with a reference not older than one year. Should the list become very long, time could be saved by listing the occurrences of the major divisions of the five fields immediately following the header. There would be one list of divisions and the addresses of their occurrences for each field. With at least 30 bytes per header and 16 bytes per list element, 100,000 directories would probably occupy 25M to 50M bytes, or about one full disk (Figure 5).

3.4 Files

It is hoped that automated records will spur reorganization of medical records with intent to save useful and only useful information. Man-machine interaction will play an important

role in this organization. However, in one experiment with nurses and automated records, nurses entered four times as many notes as normal which had half the normal useful content.

Besides deciding what is useful data, standardization is needed. As can be seen in the history or physical examination form from Mercy Hospital in Urbana, Illinois (see Appendix), the doctor is allowed a rather free hand in filling out the form. Certainly there are questions that will always be left to the discretion of the doctor, but there are several that can be answered with a yes, no, (a), (b), (c), or 98.6 in "check-off" list fashion. In addition, standard questions should be listed so that the doctor does not forget anything and the record is complete.

University of Illinois' McKinley Hospital has an automated history that is given patients periodically. The history contains 100-200 questions such as "do you smoke?" and if the answer is yes, "how many packs a day?". The answers are stored in 192 words, not compacted. When uniform standard forms are used, a "yes" answer can be stored with one bit instead of having to store an arbitrary question and answer that would consume several bytes. Standardization has other positive side effects such as making it more feasible to use para-medics for standard questioning and having the doctor add any free text notes he feels are necessary, thereby freeing the doctor from much of the routine for more important things.

Free text can be compacted about 35 per cent, depending on the character distribution. The use of variable length

character codes becomes unwieldly when an unrestricted character code such as the 88 EBCDIC character set is used. If it is found in a study of medical free text that the whole character set is not needed, then variable length character codes could be considered. Other factors affecting the decision of whether or not to use variable length code are character frequency distribution, frequency of certain combinations of characters, and the possible use of a certain subset of characters on a particular form allowing for the use of different compaction codes in different places.

One way to compact an unrestricted character set is to use the free codes for pairs of characters. For example, only 88 of a possible 256 EBCDIC codes are used, leaving 168 codes for pairs of characters. The following method employing this idea has been suggested as being efficient in compaction and expansion times. EBCDIC spreads its 88 codes over the 256 possible values. The compaction code compresses the 88 codes into the first 88 values (i.e., A is 1, B is 2, etc.). Then master and combining characters are chosen such that the number of master characters times the number of combining characters is less than or equal to 168 and such that the frequency of the combined pairs is maximized for optimal compaction. In Figure 6, eight master characters were chosen. Of the eight letters chosen, six occurred most frequently while I and U were chosen for their key position in syllables. For combining characters the 21 most frequent characters were chosen.

Master Characters		Combining Characters		Noncombining Characters						Combined Pairs			
Symbol	Base Value	Symbol	Hex Code	Symbol	Hex Code	Symbol	Hex Code	Symbol	Hex Code	Symbol	Hex Code	Symbol	Hex Code
b̸	58	b̸	ØØ	J	15	q	2B	<	41	b̸b	58	Ab̸	6D
A	6D	A	Ø1	K	16	r	2C	(	42	b̸A	59	AA	6E
E	82	B	Ø2	Q	17	s	2D	+	43	b̸B	5A	AB	6F
I	97	C	Ø3	X	18	t	2E	&	44	b̸C	5B	AC	7Ø
O	AC	D	Ø4	Y	19	u	2F	!	45	b̸D	5C	↓	↓
N	C1	E	Ø5	Z	1A	v	30	$	46	b̸E	5D	AW	81
T	D6	F	Ø6	a	1B	w	31	*	47	b̸F	5E	Eb̸	82
U	EB	G	Ø7	b	1C	x	32	)	48	b̸G	5F	EA	83
		H	Ø8	c	1D	y	33	;	49	b̸H	6Ø	↓	↓
		I	Ø9	d	1E	z	34	-	4A	b̸I	61	EW	96
		L	ØA	e	1F	Ø	35	/	4B	b̸L	62	Ib̸	97
		M	ØB	f	20	1	36	,	4C	b̸M	63	↓	↓
		N	ØC	g	21	2	37	%	4D	b̸N	64	Ob	AC
		O	ØD	h	22	3	38	_	4E	b̸O	65	↓	↓
		P	ØE	i	23	4	39	>	4F	b̸P	66	Nb̸	C1
		R	ØF	j	24	5	3A	?	50	b̸R	67	↓	↓
		S	10	k	25	6	3B	:	51	b̸S	68	Tb̸	D6
		T	11	l	26	7	3C	#	52	b̸T	69	↓	↓
		U	12	m	27	8	3D	@	53	b̸U	6A	Ub̸	EB
		V	13	n	28	9	3E	'	54	b̸V	6B	↓	↓
		W	14	o	29	,	3F	=	55	b̸W	6C	VW	FF
				p	2A	.	40	"	56				
								≤	57				

NOTE: The symbol b̸ denotes a space.

Figure 6. Compacted code.

The following is the algorithm for compaction.

Step 1. Translate EBCDIC stream into compressed code.

Step 2. Check for master characters.
If it is a master character, go to step 3.
If it is not a master character, store as is and go to step 2 for next character.

Step 3. Check next character for combining character.
If it is a combining character, go to step 4.
If it is not a combining character, store both as is and go to step 2 for next character.

Step 4. Add base value of master character to combining character and store.
Go to step 2 for next character.

For expansion, a character is comparable to 58 hexadecimal. If the character is less than 58, then a table lookup retrieves one EBCDIC character. If the character is greater than or equal to 58, then the character is multiplied by two to retrieve two EBCDIC characters from the table. On a IBM 360/40 compaction took 73 msec per 1000 characters, while expansion took 65 msec per 1000 characters of output stream.

In a discussion with a nurse in charge of hospital records, it was found that the average patient folder contained about 20 pages. Most forms the doctor had scrawled on were typed before being put into the patient folder. This indicates that since most doctors restrain themselves from writing very much and most doctors do not write as small as typing, and the forms contain sufficient space, those forms that contain text contain only about 250 words. As can be seen in the medical

chart (Figure 7), about one-third of the forms are mostly free text. At an average of 5 characters per word, the hospital file would contain 10,000 characters of free text and 5,000 characters in highly compressable information before compaction. After compaction this would reduce to 6,500 bytes of free text and 1,500 of other compacted information. This would indicate about 8,000 bytes for a hospital file of a single patient. This file includes more than one hospital stay by the patient. For the purpose of the data base, each hospital stay would have to be stored separately so that the SNOMed codes associated with each visit can address only that visit. When one considers that Mercy is only one of three hospitals serving the patient in the Champaign-Urbana area, and that Mercy serves 9,000 patients a year, most of the 100,000 patients will visit the hospital at one time or another. Even if only 50,000 patients ever visit the hospital, that would require 400M bytes just to store the hospital files.

Quite a few complex problems occur in compacting information in a hospital file as is indicated by storage of test results. Mercy Hospital, which has no real specialty, is capable of giving 243 different tests, although 27 of these account for 85 per cent of those actually given. The average patient receives about 13 tests, one of which is usually a SMA 12/60, which is 12 tests in one. With so few results (24, one for each test plus 12 for the SMA 12/60) for each patient and the fact that a patient can be given the same test more than once, it appears more efficient to just allow two bytes per test result,

1. Discharge Summary Sheet
2. Admitting Notice
3. Emergency Room Record
4. History
5. Physical Examination
6. Consultation
7. Laboratory Sheets (Date order)
   Authorization for Blood Transfusions
   Diabetic Sheet
8. X-ray Reports (Date order)
9. Electrocardiograms (Date order)
10. Electroencephalograms (Date order)
11. Physical Therapy Reports
12. Social Service Reports
13. Inhalation Therapy Reports
14. Doctors Orders and Progress Notes (Date order)
15. Graphic Charts
16. Medication Sheets
17. Intake and Output
18. Vital Signs (Date order)
19. Recertifications for Medicare
20. Release for Information Forms
21. Admission Check Off List
22. Nurses Notes (In Date order)
23. Death Certificate
24. Autopsy Report

Figure 7. Medical chart.

1. Discharge Summary Sheet
2. Admitting Notice
3. Emergency Room Record
4. History
5. Physical Examination
6. Consultation
7. Laboratory Sheets (Date order)
   Authorization for B.T.
   Diabetic Sheet
8. X-Ray Reports (Date order)
9. Electrocardiograms (Date order)
10. Electroencephalograms (Date order)
11. Physical Therapy Reports
12. Social Service Reports
13. Inhalation Therapy Reports
14. Authority to Operate
15. Anesthesia Record
16. Report of Operation
17. Tissue Report
18. Recovery Room Record
19. Doctors Orders and Progress Notes (Date order)
20. Graphic Charts
21. Medication Sheets
22. Intake and Output
23. Vital Signs
24. Recertifications for Medicare
25. Release for Information
26. Admission Check Off List
27. Nurses Notes (Date order)
28. Death Certificate
29. Autopsy Report

Figure 8. Surgical chart.

1-18. Same as Surgical Chart

19. Cardiac-Pulmonary Pump Records

20. Physical Therapy Records

21. Occupational Therapy Records

22. Inhalation Therapy Records

23. Social Service Records

24. Doctors' Orders and Progress Notes

25. Graphic Charts

26. Medication Sheets

27. Intake and Output

28. Vital Signs

29. Recertifications for Medicare

30. Patient's Transfer

31. Admission Check Off List

32. Release of Information Form

33. Nurses' Notes (Date order)

34. Death Certificate

35. Autopsy Report

Figure 9. Cardiac surgical chart.

SAME AS A MEDICAL CHART

Reports for Voluntary and Emergency Admission go on back of chart.

Figure 10. Psychiatric chart.

1. Discharge Summary Sheet
2. Admitting Notice
3. Prenatal Record
4. Release to be in Delivery Room
5. Labor-Delivery Record
6. Labor Summary
7. Laboratory Sheets (Date order)
8. X-Ray Reports (Date order)
9. Doctors' Orders and Progress Notes
10. Graphic Charts
11. Medication Sheets
12. Nurses' Notes (Date order)

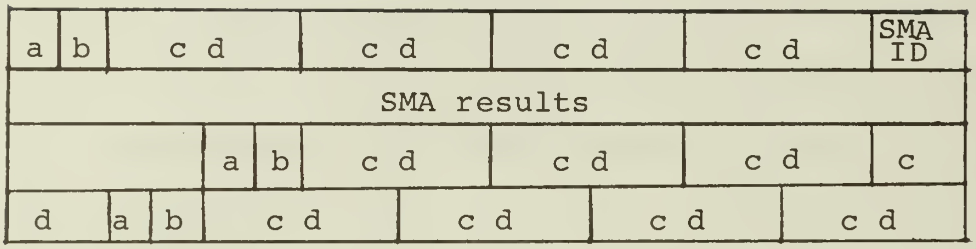
Figure 11. Obstetrical chart.

1. Discharge Summary Sheet
2. Admitting Notice
3. Newborn Physical
4. Release for Lying-In Service
5. Laboratory Sheets (Date order)
6. Report of Operation (Circumcision)
7. Doctors' Orders and Progress Notes
8. Nurses' Notes (Date order)
9. Birth Certificate

Figure 12. Newborn chart.

one for test identification and one for the result. Too much overhead would be involved in trying to use some combination of fixed area for the most frequent tests and a bit map for the rest to save the one byte for test identification, not to mention the trouble with coding two results for the same test.

Once one byte has been decided on for the result, the question arises as to whether this provides sufficient accuracy or not. This question cannot be answered with a "yes, most of the time," but must be answered, "yes, all of the time." For instance, a platelet count ranges from 0 to upward of 500,000 with a normal range for adults in the area of 200,000 to 400,000. The accuracy, though, is only two significant decimal digits. Hence, the result can be stored as a number between 0 and 256 and then multiplied by 10,000 on retrieval. For slightly more accuracy at the expense of range, the result can be stored as a multiple of 5,000. Some tests may require three significant decimal digits. The test for sodium, for example, is sensitive and requires three decimal digits of accuracy, but the normal range for this test is 136-150 mEq./XL. which is right in the middle of the one byte 0-256 range. Therefore, the sodium test will have three decimal digits of significance with no modification to force a fit into the one byte range. In conclusion, one byte will always allow two significant decimal digits and sometimes three, depending on the standard deviation and the mean of the result. Since the necessary accuracy for each test is erratic, a test by test study will have to be made concerning accuracy needed and available for each test. The hospital file and



a) Difference in days between test data and admittance date (4 bits)

b) Number of test given that day (4 bits)

c) Test identification number (1 byte)

d) Test result (1 byte)

Figure 13. Laboratory tests.

private doctors' file will be the most referenced files for medical care. Other files of importance include the family, social, and personality histories and the private doctors' reports. The family history would contain history relating to the patient or, to say it another way, family history not found in the files of the patient's spouse, children, parents, brothers, or sisters. It would include the patient number of parents, spouse, children, brothers, or sisters. Other items included would be the history of the income of the patient and his past jobs, homes (size, number of rooms, living conditions), types of neighborhoods, access to playgrounds, etc. Also a list of diseases that have occurred in the family and are important, such as tuberculosis, allergies, diabetes, etc., would be included in the file. Immunization records (type, number, age, and reactions) and serious illnesses of the patient would be listed.

In the social history, the patient's history of interaction with the public would be listed. This would include a history of schools (public or private, overcrowded, type of students, class, grades, nursery school, special aptitudes), clubs (political, professional, school or recreational), church affiliation and type of home (country house, suburban house, duplex, apartment or high-rise apartments).

The personality history would include the patient's relations with others and any history that would affect these relations. A history of childhood development and nutrition up to about the first year would be included. These records would

be the routine check-ups given by the doctor. Also included would be habits (eating, sleeping, exercise, etc.), attitudes toward school, and relations to others. Relations between patient and mother, father, spouse, and children would note any shyness, submissiveness, separation, and negativism. Notice of extraordinary relations with any hobbies would also be included, as would unusual relations with others. In addition, physical deformities affecting the patient's personality would be listed.

The histories are prime candidates for automation. The patient could answer periodic updating questionnaires for family and social histories while waiting for a doctor's appointment. The personality history would need updating from both the patient and the doctor. Assuming highly compressable data, 2000 bytes should be sufficient for each history. This, though, would require an additional 600M bytes of storage.

The private doctor's report would differ somewhat from a hospital file, since a hospital collects a large amount of information in a short period of time, a few days. On the other hand, a doctor might see a patient for a half-hour every month or every six months, or just once. A dermatologist, for example, might prescribe to a patient a medication for a skin infection and instruct the patient not to return if the infection goes away. The doctor's entry into CMHDB might only contain the SNOMed code, date, prescription, and attending doctor. Rather than add four bytes of reference to address three, a miscellaneous file could be established that was searched by date. The

entry in the directory would consist of the SNOMed code, number of reference bytes, date, and one byte pointing to the address of the miscellaneous file. If noticeable activity occurred in any one area, say the number of references reached a certain number for a particular code, a separate file would be formed containing all entries for that code. These entries would then be eliminated from the miscellaneous file. Doctors, such as dentist and optometrist, would most frequently use the miscellaneous file while those more dependent on the past history and present status of the patient, such as general practitioner, would have a separate file for that patient. The average storage requirements for a private doctor's reports are expected to be 1,000 bytes of compacted information per patient.

CHAPTER 4

CONCLUSION

As has been pointed out, the technology is available to build a CMHDB. Legal problems and the traditional way of doing things will inhibit the feasibility and efficiency of a CMHDB. In particular, the system must be used as a primary source of information. It will not be economically feasible or more than a research toy as long as duplicate records are manually recorded. Secondly, the efficiency of the system will be greatly enhanced if problem oriented records and SNOMed type codes are adopted. There is a trend toward problem-oriented records, but it is difficult to change habits 100 years old. Problem-oriented records allow easy search and retrieval while making it possible to restrict user access.

A centralized data base is reduced, first, by minimization of redundancy, and secondly, by data compaction techniques. Centralization offers an economically viable method of using the best equipment for information storage and retrieval. Finally and most importantly, a Community Medical History Data Base offers the most complete records for more individualized care. Such a system would be ideal for a Health Maintenance Organization (HMO). See, in particular, [53], where dentistry, psychiatry, and general medical care are all practiced in one building.

The system discussed was required to be on-line with a response that would satisfy an impatient doctor. The core of

the desired computer must contain programs, translation tables, organization charts, user directory, and 300K bytes for the table of patient directory addresses. This would indicate a need of a computer with a .75 to 1.0 million bytes of core or primary storage. The computer also needs to be able to handle up to 100 requests at once. This is based on the assumption that private doctors would make one request every half-hour and since most private practitioners keep office hours, there would be a large influx of requests in a short period of time about 8:30 a.m. The patient directories would require a disk the size of the 69.8M byte IBM M3340. About 8 IBM 2314 Disk Array units would be needed for the medical files. The eight disk array units would allow about 1.5 billion bytes of storage. If this was thought to be insufficient, consideration to magnetic domain (bubble) memories, charge-couple devices or optical memories should be given. All are reputed to be faster, larger, and less expensive than disk and practical within two or three years. Such a system could be implemented for 1.5 million dollars annually, including overhead. This is $3,000 per user, but cost would be proportional according to use. Most of the benefits to the private doctor would be intangible. Records would be more complete and readily available. An increase in speed and accuracy in information retrieval and decrease in paper and possibly staff could be achieved. Medical records would be consolidated for more effective and efficient use. The complete records allow the evaluation of care rendered and the review of utilization of health facilities, medical care,

and services. Cost to the users could be reduced by charging statisticians, researchers, and educational and training centers for use of CMHDB.

It is hoped this provides some insights into the problems and solutions of designing a community medical history data base.

APPENDIX

TESTS GIVEN AT MERCY HOSPITAL IN URBANA, ILLINOIS

TESTS	FREQUENCY PER 1000 TESTS
SURGICAL PATHOLOGY	
Autopsies:	
Coroner's Case	*
Mercy	*
Private Cases	*
Stillborn	*
Tissues:	
Bone Marrow	*
Frozen Section	1
Stone Analysis	*
Gross	24
Microscopic	21
Pap Smear	20
Programmed Medical Histories	3
CLINICAL PATHOLOGY	
Blood Bank:	
Blood Transfusions	14
Transfusion Reactions	*
Crossmatch	26
Type	49
Rh. Factor	49
Rh. Titer	*
Coombs	*
Cord Blood Studies	*
Selectogen	11
Rhogam	*
Phlebotomy	*
Cryoprecipitates	3
Platelet Concentration	*
Spinal Fluid:	
Cell Count	2
Chloride	1
Colloidal Gold	*
Glucose	1
Pandy	*
Protein	2
VDRL	2
Spinal Electrophoresis	*
LDH Electrophoresis	*
Spinal Viral Antigens	*
Spinal Isoenzymes	*

TESTS	FREQUENCY PER 1000 TESTS
Biochemistry:	
Acetone	*
Albumin	*
Ammonia	*
Amylase	4
Blood Alcohol	*
BSP	*
Bilirubin (Direct & Indirect)	4
Bilirubin (Total Only)	*
Calcium	*
Carotene	*
Ceph. Floc.	*
Copper	*
CPK	7
Cholesterol	*
Creatinine	3
Creatinine Clearance	*
D-Xylose Tolerance	*
LDH	4
Gastric Analysis	*
Glucose	27
Glucose Tolerance	1
Icterus Index	*
Insulin Tolerance	*
Iodine T3	3
Iodine T4	3
Iodine PBI	1
Free Thyroxine Index	3
Iron	*
Iron Binding Capacity	*
Iron Saturation	*
Lactose Tolerance Test	*
LAP	*
Lipase	*
Magnesium	*
NPN	*
Phosphatase (Acid)	*
Phosphatase (Alkaline)	*
Phosphorous	*
Oxygen Saturation	*
Total Protein	*
SMA 12/60	71
Sweat Test	*
Sweat Electrolytes	*
Thymol Turbidity	*
SGOT	5
SGPT	*
Urea Nitrogen (BUN)	6

TESTS	FREQUENCY PER 1000 TESTS
Biochemistry (continued):	
Uric Acid	*
Zinc	*
G6PD	*
Lithium	1
ALA-PDG	*
Australian Antigen	*
Sodium	21
Potassium	23
Chloride	21
Base Excess	15
Hydrogen pH	15
p02	15
pC02	15
Oxygen Content	*
HEMATOLOGY:	
CBC	95
Hematocrit	18
Hemoglobin	13
WBC	5
Differential	2
Bleeding Time	*
Blood Volume	*
Clotting Time	3
Clot Retraction	*
Color Index	*
Eosinophil Count	*
Fibrinogen (Fibrindex)	*
Fibrinolysin	*
LE Prep	*
Methemoglobin	*
PTT	12
Platelet Count	5
Reticulocyte Count	*
RBC Count	*
RBC Indices	*
RBC Fragility	*
Sedimentation Rate	9
Sickle Cell Prep	3
Prothrombin Time	20
Peripheral Smear	*
RBC Study for Basophil Stipp.	*
Platelet Concentration	*
Sperm Count	*
Protamine	*

TESTS	FREQUENCY PER 1000 TESTS
MICROBIOLOGY:	
Bacterial Culture	30
AFB Culture	3
Fungi Culture	2
Sensitivity	17
Smear, Bacterial	1
AFB	1
Fungi	1
Fluorescent Microscopy	*
PKU	6
Viral Studies	*
SEROLOGY:	
Cold Agglutination	*
Febrile Agglutination	*
Antinuclear Antibody	*
ASO Titer	*
Brucella	*
C-Reactive Protein	*
Gamma Globulin	*
Heterophile	*
Histoplasmosis	*
RA Latex	*
Rubella	*
Pregnancy Test	*
Tularemia	*
VDRL	62
FTA	*
Plasma **	*
Darkfield	*
URINALYSIS:	
Routine Ua	85
Microscopic Only	*
Hemoglobin	*
Bile	*
Ketone Bodies	*
Glucose	*
pH	*
Bence Jones Protein	*
Protein	*
Albumin	*
PSP	*
Specific Gravity	*
Urobilinogen	*
Porphyrins	*
Diacetic Acid	*
Urine Lactose	*
Addis Count	*

TESTS	FREQUENCY PER 1000 TESTS
FECES:	
Bile	*
Fat	*
Occult Blood	2
Ova and Parasites	1
Starch	*
Trypsin	*
Eosinophiles	*
HORMONES:	
Aldolase	*
Aldosterone	*
Catecholamines	*
Estrogen	*
Gonadotropin	*
5HIAA	*
17-Corticosteroids	*
17-Hydroxycorticosteroids	*
17-Ketosteroids	*
17-Ketogenic Steroids	*
Porphobilinogen	*
Porphyrins	*
Pregnanediol	*
Pregnanetriol	*
Sertonin	*
Urobilinogen	*
VMA	*
FSH	*
Testosterone	*
Haptoglobins	*
Masters EKG	*
FUNCTIONAL STUDIES:	
BMR	*
EEG	2
EKG	32
10-hr EKG	*
Pulmonary Function	*
SPECIAL CHEMISTRY:	
Serum Protein Electrophoresis	6
LDH Electrophoresis	*
Hgb Electrophoresis	*
Immunoglobulins	*
Lipids	*
Lipoprotein Phenotyping	*
Bromide Levels	*
Eatonagent Comp. Fixation	*
Immunoglobulin Electrophoresis	*

TESTS	FREQUENCY PER 1000 TESTS
SPECIAL CHEMISTRY (continued):	
Plasma Protein Electrophoresis	*
Nitratenylsis	*
Protein Ox	*
Chromosomes	*
Buccal Smear	*
Estriol	*
Blood Renins	*
Barbiturates	*
Salicylates	*
Schilling Test	*
Ewal	*
Viral Antigen Test	*
Sulfa Level	*
Triglycerides	*
Amino Acids	*
Urine for Lead	*
Osmolarity	*
Haptoglobin	*
Coma Screen	*
Vitamin A	*
Amniocentesis Fluid	*
Etholchloronol	*
Lutenizing Hormone	*
Lead	*
Plasma Cortisol	*
Beta C Fixation	*
Delta Levine	*
Folic Acid	*
Chromium 51	*
Euglobulin	*
Factor VIII	*
I3AA	*

Note: Frequency is based on 102,072 tests given over a ten month period.

* indicates a frequency of less than 1 per 1000 tests.

MERCY HOSPITAL, Urbana, Illinois

Date	DOCTOR'S ORDERS	Date	PROGRESS RECORD

DOCTOR'S ORDERS-PROGRESS RECORD

Code 5119

MERCY HOSPITAL

Patient's Name ____________________

AGE ON ADMISSION
M ○ Months
D ○ Days
O ○ Years plus 100
N ○ Newborn
Y ○ Years

SEX
○ M Male
○ F Female

ORIGIN
○ W White
○ N Negro
○ O Other

PATIENT NUMBER

ADMISSION DATE
Day Month Year

Days of Week
○ 1 Monday
○ 2 Tuesday
○ 3 Wednesday
○ 4 Thursday
○ 5 Friday
○ 6 Saturday
○ 7 Sunday

DISCHARGE STATUS
○ 1. With Approval
○ 2. Against Advice
○ 3. Transferred
○ 4. Unimproved
○ 5. Died over 48 Hours
○ 6. Died under 48 Hours
○ 7. Autopsy
○ 8. No Autopsy

HOSPITAL SERVICE
○ 1 Medical
○ 2 Surgery
○ 3 Pediatrics
○ 4 Psychiatric
○ 5 OB
○ 6 NB
○ 7 UROL
○ 8 GYN
○ 9 Orthopedics
○ 10 Ophthalmology
○ 11 Otolaryngology
○ 12 Neurosurgery
○ 13 Thoracic Surgery

DISCHARGE DATE
DAY MONTH YEAR

PAYMENT
○ 1. Blue Cross
○ 2. Commercial Ins.
○ 3. Workmans Comp.
○ 4. IPAC, ADC
○ 5. Medicare
○ 6. Private Pay
○ 7. Gov't Agencies

HEMATOLOGY
○ 1 BLEEDING COAG TIME
○ 2 CBC
○ 3 BONE MARROW

BLOOD TYPE
○ 1- O+
○ 2- A+
○ 3- B+
○ 4- AB+
○ 5- O-
○ 6- A-
○ 7- B-
○ 8- AB-
○ 9- RH-
○ A- RH +

URINALYSIS
○ 1 - Done
○ 2 - Done Later

SEROLOGY
○ 1-VDRL - Po

HEMOGLOBIN or HEMATOCRIT

WHITE BLOOD COUNT

NUMBER OF TRANSFUSIONS

TEMPERATURE

BLOOD PRESSURE

FINAL DIAGNOSIS

Pathology Report In Chart
- 1 Gross
- 2 Microsiopic
- 3 F.S.
- 4 Pap Smear
- 5 Tiss. Removed no report
- 6 No Tiss. Removed

THERAPY
- 1 X-Ray, RX
- 2 Speech
- 3 Physical
- 4 Shock
- 5 Occup.
- 6 Inhal.

PATIENT NUMBER

ATTENDING PHYSICIAN

X - RAY
- 1 Chest
- 2 Other Resp.
- 3 Skeletal
- 4 Digestive Tract & GB
- 5 Urogenital
- 6 Cardiovascular
- 7 CNS - CNS Spaces
- 8 Mammograms
- 9 Pleural & Peritoneal
- A Brain Scan
- B Liver Scan

BACTERIOLOGY
- 1 Bacterial Smear
- 2 Culture
- 3 Parasitology

BODY FLUIDS
- 1 Spinal
- 2 Other Body Fluids

FUNCTIONS
- 1 EKG
- 2 BMR
- 3 Pulmonary
- 4 EEG

TIME OF FIRST SURGERY

If began within 6 hours of admission mark here

Otherwise number of Days Preop. Stay

CONSULTING PHYSICIANS

OPERATIONS

OPERATIONS

PATIENT NUMBER

OPERATING SURGEON

SECOND SURGEON

SURGEON

OPERATIONS

DISCHARGE SUMMARY SHEET
Mercy Hospital - Urbana, Illinois

Family Name	Room No.	Hosp. No.
Attending Physician	Date of Admission	Date of Discharge

CODE NUMBER

Provisional Diagnosis

Final Diagnosis

Operation:

Brief History and Essential Physical Findings:

Significant Laboratory, X-ray, and Consultation Findings:

Course in Hospital with Complications, .·ıy

Condition, Treatment, Final Disposition on Discharge and Prognosis:

Date: ____________________

I have Examined and Approved this Completed Record

Signed ________________________________ M.D.

MERCY HOSPITAL, Urbana, Illinois

M. D.

HISTORY

MERCY HOSPITAL, Urbana, Illinois

ON THIS EXAMINATION, THE FOLLOWING SHOULD BE INCLUDED:
Skin; head--eyes, ears, nose, mouth;throat; neck; lymphnodes; chest; breasts; heart; lungs; blood vessels; abdomen, genitalia; rectal; bones--joints, muscles; extremities; neurological.

Provisional Diagnosis: ______________________________

______________________ M D

PHYSICAL EXAMINATION

MERCY HOSPITAL, Urbana, Illinois

From: Attending Physician | To: Consultant

Findings:

Diagnosis:

Recommendations:

Date ____________________ ____________________ M.D.

Signature of Consultant

Code 5107

CONSULTATION

MERCY HOSPITAL, Urbana, Illinois

Surgeon | Assistant | Date

Preoperative Diagnosis

Postoperative Diagnosis

Operation

Findings: (including the condition of all organs examined) and Procedures (including incision, ligatures, sutures, drainage, sponge count and

Wound primarily clean

Wound primarily infected

Healing of wound: Clean--primary intention

Granulations:

Stitch abscess:

Hematoma:

Deep sepsis:

____________________ M.D.

CODE 8204

REPORT OF OPERATION

LIST OF REFERENCES

[1] Amarel, S., and Kulikowski, C. "Medical Decision Making and Computer Modeling." Dept. of Computer Science, Rutgers University, New Brunswick, New Jersey, Technical Report, 2 pages.

[2] ______. "Are Office Doctors Collecting Enough Data?". Medical World News, Jan. 14, 1972, pp. 18-20.

[3] ______. Control Data Medlab Systems. Control Data Corp., Analog-Digital Systems Division, La Jolla, California, General Technical Proposal Q0049-1-RJB, Publication No. X0010323, Jan. 1970.

[4] ______. "The Medical History," Feasibility Study of Automated Medical Examining System, Philco Ford Corp., Palo Alto, Calif., Section 4, Sept. 1970, 16 pages.

[5] ______. "Multiphasic Testing." Medical World News, Oct. 15, 1972, pp. 51-62.

[6] ______. "Right of Privacy and Medical Computing." Datamation, Vol. 16, No. 4, April 1970, pp. 173-178.

[7] Balintrg, J. L. "The Camp System for Computerized Menu Plans." Hospitals, J.A.H.A., Vol. 45, May 16, 1971, pp. 92-96.

[8] Ball, M. "Effectiveness of On-Line Data Acquisition in the Clinical Laboratory," Journal of the Assoc. for the Advancement of Medical Instrumentation, Vol. 6, No. 1, Jan.-Feb. 1972, pp. 28-31.

[9] Barnett, G. O., Greenes, R. A., and Grossman, J. H. "Computer Processing of Medical Test Information." Methods of Information in Medicine, Oct. 1969, pp. 177-182.

[10] Barnett, G. O., and Greenes, R. A. "Interface Aspects of A Hospital Information System." Annals of the New York Academy of Sciences, Sept. 30, 1969, pp. 756-768.

[11] Baruch, J. J., and Barnett, G. O. "Joint Venture at Massachusetts General." Datamation, Vol. 11, No. 12, Dec. 1965, pp. 29-33.

[12] Berlin, M. M. "Computers in the Laboratory." Computers and Automation, Vol. 19, No. 6, June 1970, pp. 24-27.

[13] Bond, V. B., Bowman, C. M., Lee, N. L., Petersen, D. R., and Reslock, M. H. "Interactive Searching of a Structure and Biological Activity File." Journal of Chemical Documentation, Vol. 11, No. 3, 1971, pp. 168-170.

[14] Cashman, M. W. "A Read/Write Optical Memory System." Datamation, Vol. 19, No. 3, March 1973, pp. 66-69.

[15] Clary, M. D. "Data Processing Requirements for Health Insurance." Datamation, Vol. 16, No. 3, March 1970, pp. 81-86.

[16] Coe, F. L. "A Computer Based Diagnostic and Data Management Program for Kidney Stones." Michael Reese Hosp. and Med. Center, Pritzker School of Medicine, Univ. of Chicago, Chicago, Illinois, 3 pages.

[17] Craig, J. L., and Derryberry, O. M. "Applied Concepts of Automation in an Occupational Medical Program." Industrial Medicine, Vol. 40, No. 5, August 1971, pp. 9-17.

[18] Dearden, J., and McFarlan, F. W. Management Information Systems: Text and Cases, Richard D. Irwin, Inc., Homewood, Illinois, 1966.

[19] Gabrielli, E. R. "Dual Data System in the Hospital." Journal of Clinical Computing, Dec. 1971, pp. 4-7.

[20] Gabrielli, E. R. "Organization of the Information Contained in the Clinical Chart." Journal of Clinical Computing, June 1971, pp. 1-7.

[21] Gabrielli, E. R., Newcombe, H., Teitelbaum, L. E., Shepley, M. D., and Sagen, O. K. Blueprint of the Western New York Health Data Network, Report of the Medical Society of the County of Erie, New York.

[22] Geisler, R. "The Thomis Medical Information System." Datamation, Vol. 16, No. 6, June 1970, pp. 133-136.

[23] Gouvela, W. A., Diamantis, C., and Barnett, G. O. "Computer Applications in the Hospital Medication System." American Journal of Hospital Pharmacy, Mar. 1969, pp. 140-150.

[24] Greenes, R. A. "Medical Information Systems: A Look Ahead." Conf. on Computers in Radiology, Columbia, Missouri, Sept. 1970, 14 pages.

[25] Greenes, R. A., Barnett, G. O., Klein, S. W., Robbins, A., and Prior, R. E. "Recording, Retrieval and Review of Medical Data by Physician Computer Interaction." New England Journal of Medicine, Feb. 5, 1970, pp. 307-315.

[26] Greenes, R. A., Pappalardo, A. N., Marble, C. W., and Barnett, G. O. "A System for Clinical Data Management." Proc. of the Fall Joint Comp. Conf., Vol. 35, 1969, pp. 297-305.

[27] Haggerty, J. R. "Computerized Information System." Hospitals, J.A.H.A., Vol. 44, Nov. 1, 1970, pp. 43-46.

[28] Hall, P., Mellner, C. H., and Danielsson, T. "A Data Processing System for Medical Information," Methods of Information in Medicine, Vol. 6, No. 1, Jan. 1967, pp. 1-6.

[29] Hemmer, J. "The Data Transmission System." Upstate Medical Center, Syracuse, New York, Technical Report, 7 pages.

[30] Hoffman, P. B., Grossman, J. H., Thoren, V. J., and Barnett, G. O. "Automated Patient Census Operation: Design, Development, Evaluation," Hospital Topics, May, 1969.

[31] Holmgren, J. H., and Zimmerman, W. "Computerized Ordering: How it Works, What it Does." The Modern Hospital, March 1971, p. 70.

[32] Hopp, D. I., and Cable, J. D. "The Construction of the Clinical Data Base." Journal of the Assoc. for the Advancement of Medical Instrumentation, Vol. 6, No. 1, Jan.-Feb. 1972, pp. 32-37.

[33] Jackson, G. G. "Information Handling Costs in Hospitals." Datamation, Vol. 15, No. 5, May 1969, pp. 56-59.

[34] Jackson, G. G. "The Role of Administrators and Physicians in the Development of Hospital Information Systems." Computers and Automation, Vol. 19, No. 6, June 1970, pp. 33-35.

[35] Jacobs, J. F. Jr., Davis, R. F., and Bakerman, S. "Computer Diagnosis Generated from SMA-12/60 Chemistry Profiles." Journal of the Assoc. for the Advancement of Medical Instrumentation, Vol. 6, No. 1, Jan.-Feb. 1972, pp. 37-42.

[36] Jenkins, D. R. "Problems of Computer Applications in Medical Research." Transactions New York Academy of Sciences, 1966, pp. 439-447.

[37] Josephus, M., and Venarda, M. "How Mercy Planned its Communications." The Modern Hospital, August 1966, pp. 87-89.

[38] Jydstrup, R. A., and Gross, M. J. "Cost of Information Handling in Hospitals." Health Services Research, Winter 1966, pp. 235-271.

[39] Katona, P. G., Pappalardo, A. N., Marble, C. W., Barnett G. O., and Pashby, M. M. "Automated Chemistry Laboratory: Application of a Novel Time-Shared Computer System." Proc. of the IEEE, Vol. 57, No. 11, Nov. 1969, pp. 2000-2006.

[40] Liu, H. "A File Management System for a Large Corporate Information System Data Bank." Proc. of the Fall Joint Comp. Conf., Vol. 13, 1968, pp. 145-156.

[41] Matthews, D. Q. The Design of the Management Information System. Averbach Publishers, New York, New York, 1971.

[42] McLaughlin, R. A. "A Bleak Future for Discs?" Datamation, Vol. 19, No. 4, April 1973, pp. 106-108.

[43] McNabb, M. E. "90-Day Nonselective Menus by Computer." Hospitals, J.A.H.A., Vol. 45, May 16, 1971, pp. 88-91.

[44] Mueller, W. J., Walsh, L. F., Hemmer, J. A., and Allen, F. H. "Patient Oriented Hospital Information System." Workshop on Hospital Information Systems 23rd ACEMB, Washington, D.C., November 1970, 5 pages.

[45] Mulford, J. E., and Ridell, R. K. "Data Compression Techniques for Economic Processing of Large Commercial Files." ACM Symposium on Information Storage and Retrieval, 1971, pp. 207-215.

[46] Myers, J., Gelblat, M., and Enterline, H. T. "Automatic Encoding of Pathology Data." Archives of Pathology, Vol. 89, Jan. 1970, pp. 73-78.

[47] O'Brien, J. J. Management Information Systems Concepts, Techniques, and Applications, Van Nostrand Reinhold Co., New York, 1970.

[48] Opit, J. J., and Woodroffe, F. J. "Computer-Held Clinical Record System--I, Description of System." British Medical Journal, Vol. 4, Oct. 1970, pp. 76-79, and "Computer-Held Clinical Record System--II, Assessment." British Medical Journal, Vol. 4, 1970, pp. 80-92.

[49] Pendergrass, H. P., Greenes, R. A., Barnett, G. O., Poitras, J. W., Pappalardo, A. N., and Marble, C. W. "An On-Line Computer Facility for Systematized Input of Radiology Reports." Radiology, March 1969, pp. 709-713.

[50] Pollycove, M. "Benefits and Disadvantages of Clinical Laboratory Computerization." Journal of the Assoc. for the Advancement of Medical Instrumentation, Vol. 6, No. 1, Jan.-Feb. 1972, pp. 42-46.

[51] Reilly, N. B. "Computers in Medicine." Datamation, Vol. 15, No. 5, May 1969, pp. 46-49.

[52] Reingold, E. M. "Notes on AVL Trees." Department of Computer Science, University of Illinois, Urbana, Illinois, Report No. 411, May 1971.

[53] Rothfeld, M. "Sensible Surgery for Swelling Medical Costs." Fortune, Vol. 86, April 1973, pp. 110ff.

[54] Ruderman, M. "The Hospital Computer Comes of Age." Computers and Automation, Vol. 19, No. 6, June 1970, pp. 28-32.

[55] Ruth, S. S., and Kreutzer, P. J. "Data Compression for Large Business Files." Datamation, Vol. 18, No. 9, Sept. 1972, pp. 62-66.

[56] Schwartz, M. D. "Status of Hospital Information Systems." Hospital Progress, June 1970, pp. 53-60.

[57] Silver, H. K., Kempe, H., and Bruyn, H. B. Handbook of Pediatrics, Lange Medical Publications, Los Altos, California, 1969.

[58] Singer, J. P. "Computer-Based Hospital Information Systems." Datamation, Vol. 15, No. 5, May 1969, pp. 38-45.

[59] Singer, J. P., and Petro, F. A. "A Case History: Implementation of a Computer-Based Patient Accounting System." Computers and Automation, Vol. 19, No. 6, June 1970, pp. 19-22.

[60] Smith, R. M. "How To Automate A Hospital." Management Services, Jul.-Aug. 1966, pp. 48-53.

[61] Snyderman, M., and Hunt, B. "The Myriad Virtues of Text Compaction." Datamation, Vol. 16, No. 16, Dec. 1970, pp. 36-40.

[62] Souder, J. J. "Computers Can Bring a New Rationality into Hospital Design." The Modern Hospital, Vol. 108, March 1968, pp. 80-86.

[63] Spann, M. L., and Willis, D. D. "A Comparative Study of a Fragmentation vs. a Topological Coding System in Chemical Substructure Searching." Journal of Chemical Documentation, Vol. 11, No. 1, Nov. 1971, pp. 43-47.

[64] Springer, E. W. Automated Medical Records and the Law, Health Law Center, Division of Aspen Systems Corporation, Pittsburgh, Penn., 1971.

[65] Springer, G. D., Jr. "One Computer Works for Three Hospitals." The Modern Hospital, Vol. 107, No. 1, August 1966.

[66] Wagner, R. A. "Common Phrases and Minimum Space Text Storage." Communications of the ACM, Vol. 16, No. 3, March 1973, pp. 148-152.

[67] Walsh, L. F. "Data Communication System for On-Line Hospital Computer Applications." State University of New York, Syracuse, New York, 11 pages.

[68] Weed, L. L. Medical Records, Medical Education, and Patient Care, The Press of Case Western Reserve Univ., Cleveland, Ohio, 1970.

[69] Wells, A. H. "The Lexicon of Medical Practice for Computer." College of American Pathologists, Sept. 1972, pp. 229-237.

[70] Wright, M. A. "Mechanizing a Large Index." The Computer Journal, Vol. 3, No. 2, July 1960, pp. 76-83.

[71] Yarnall, S. R. "The Central Role of the Data Base History in Health Screening." University of Washington School of Medicine, 15 pages.

BIBLIOGRAPHIC DATA SHEET	1. Report No. UIUCDCS-R-73-589	2.	3. Recipient's Accession No.

4. Title and Subtitle
Architecture of a Community Medical History Data Base

5. Report Date
August 7, 1973

6.

7. Author(s)
Craig Alan Mills

8. Performing Organization Rept. No. UIUCDCS-R-73-589

9. Performing Organization Name and Address
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

10. Project/Task/Work Unit No.

11. Contract/Grant No.

12. Sponsoring Organization Name and Address
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

13. Type of Report & Period Covered
Thesis Research

14.

15. Supplementary Notes

16. Abstracts
The medical profession has many cumbersome methods, such as most present medical codes, handwritten records, and the amount of paperwork involved in the care of a patient. These cumbersome methods and the establishment of Medicare, insurance forms, and other red tape which have become an intricate part of our society hinder the development of computer applications in the medical profession. This paper attempts to cover the many legal, security, and technical problems that arise. A general background of computer applications in the medical professions is given with emphasis on the usefulness of a medical history data base. From such computer applications, the medical profession will derive innumerable benefits. These benefits include more complete records, more individualized care, and efficient handling of paperwork.

17. Key Words and Document Analysis. 17a. Descriptors

Data base, information retrieval, automated medical records, data compaction, computers and the hospital, community medical record data base.

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Unlimited Release	UNCLASSIFIED	75
	20. Security Class (This Page)	22. Price
	UNCLASSIFIED	-----